LA-UR-20-27867
Approved for public release; distribution is unlimited.

Title: Validation and Testing of SALSA3D: A global model of compressional wave speed for the crust and mantle

Author(s): Begnaud, Michael Lee
Ballard, Sanford
Hipp, James R.
Encarnacao, Andre V.
Young, Christopher J.
Phillips, William Scott
Rowe, Charlotte Anne

Intended for: Report

Issued: 2021-11-22 (rev.1)

Validation and Testing of SALSA3D: A global model of compressional wave speed for the crust and mantle

Begnaud, Michael L.
Ballard, Sanford
Hipp, James
Encarnacao, Andre
Young, Christopher
Phillips, W. Scott
Rowe, Charlotte A.

Introduction

For treaty monitoring, the ability to accurately locate seismic events is essential. Since many event location techniques rely on the comparison of observed to predicted seismic wave travel times through the Earth, it is important to be able to accurately estimate travel times and their uncertainties.

Given the limited station coverage of a typical global monitoring network, such as the International Monitoring System (IMS) of the Prepatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO PrepCom), it is likely that such events will be detected by very few stations, providing poor network geometry. For locating such events, it is essential to have extremely high-fidelity travel-time predictions, particularly at regional distances where lateral heterogeneity can be significant. Accurately accounting for lateral heterogeneity implies using 3D models of the Earth to calculate travel times, but relatively few of the available 3D Earth models are appropriate for high fidelity travel-time prediction, and it is unclear whether any of those that are appropriate actually improve event locations.

The SAndia LoS Alamos 3D (SALSA3D) tomography model (version 2, or V2) was formulated and computed with the intent of estimating travel-time predictions and their corresponding uncertainties as accurately as possible (Ballard *et al.*, 2016a), specifically for improving seismic event location accuracy and precision. Using two separate validation event sets and data selection configurations, we present validation of version 2 of the SALSA3D model compared to one-dimensional (1D), two-dimensional (2D), as well as other three-dimensional (3D) seismic P-wave velocity models.

SALSA3D Velocity Model

The SALSA3D model was developed with the main goal of improving seismic location accuracy and precision when using regional (Pn) and teleseismic (P) arrivals, either solely or in combination, as these are the most common arrivals used for routine seismic location. Ballard et al. (2016a) fully describes the development and description of the SALSA3D model with variable resolution 3D ray tracing (Ballard *et al.*, 2009). Of particular importance in Ballard et al. (2016a) is the determination of the full 3D covariance matrix resulting from the inversion of body wave travel time residuals and the calculation of path-dependent travel time uncertainties (PDU) from this 3D covariance matrix. The capability to calculate specific PDU estimates is critical when using the SALSA3D model for seismic event location. The standard practice of using a simple 1D distance-dependent uncertainty (1DU) estimate with distance is not fully adequate for the highly variable 3D velocity structure resulting from the tomographic inversion. Ballard et al. (2016a) demonstrates that patterns of 3D travel time uncertainty follow data coverage in the model, with low uncertainty where ray paths are abundant and higher uncertainties where there is a paucity of ray paths (Figure 1).

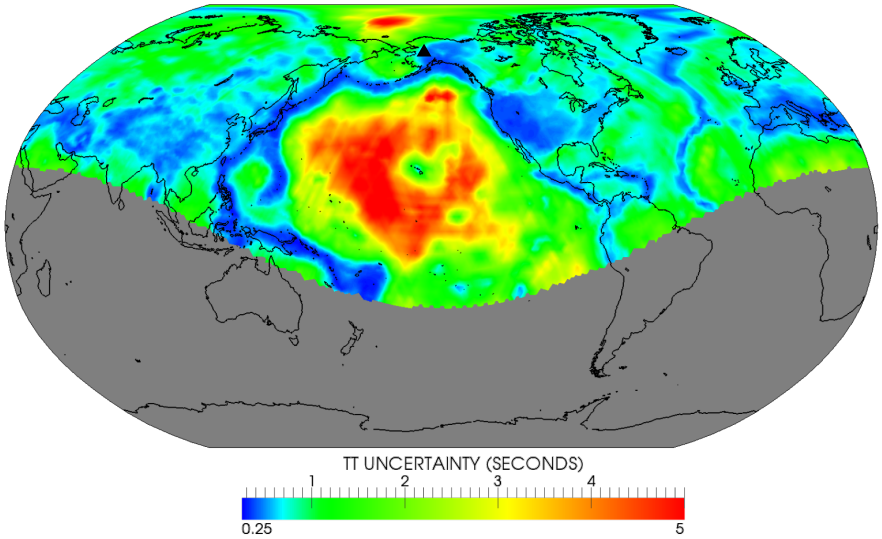


Figure 1. Travel time uncertainty calculated for station ILAR in Alaska. The pattern of uncertainty values follows the general pattern of ray path density through the model.

The SALSA3D model (version 2.1) (www.sandia.gov/salsa3d) is developed and stored as a triangular tessellation using the GeoTess open-source software (Ballard *et al.*, 2016b) (www.sandia.gov/geotess), with approximately 1 degree node separation and built-in ellipticity (i.e., WGS80) such that no ellipticity corrections are necessary when using this type of model (Ballard *et al.*, 2009). The model has variable grid spacing, both laterally and with depth, and triangle grid size related to local ray path density. While recent updates to the SALSA3D model include both shear-wave velocity (version 2.1ps) as well as core phases and underside reflection phases (version 3), this paper describes the validation of the original released version of the SALSA3D model, using first-arriving P phases (i.e., regional Pn and teleseismic P).

As a first look into validation of the SALSA3D model, we plot the average and standard deviation of travel time residuals with distance through the starting model (before tomography) and the final 3D model (after tomography) (Figure 2). After tomography, the SALSA3D model shows significantly reduced average travel time residuals as well as travel time residual spread with distance. The final root-mean squared (RMS) residual is 0.94 s, a reduction of 50% from the starting model.

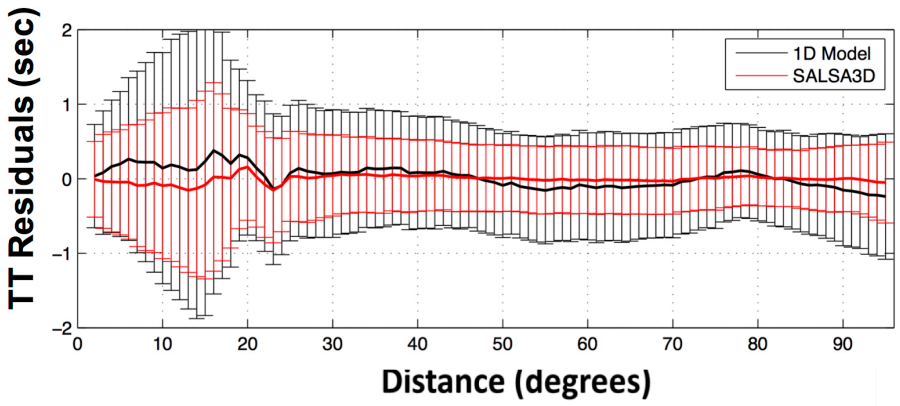


Figure 2. Travel time residuals with distance shown as an average (line) and standard deviation (error bars) for the starting model (black) and final SALSA3D model after tomography (red).

Metrics and Validation Events

Metrics

For any model validation, we are interested in several metrics to indicate model performance with regard to location accuracy and precision. Depending on the validation event set, events will be relocated with the *ak135* model and its associated uncertainty (Kennett *et al.*, 1995), the updated Regional Seismic Travel Time (RSTT) model version with PDU ("pdu2020012Du", www.sandia.gov/rstt) (Begnaud *et al.*, 2020a; Begnaud *et al.*, 2020b), the Source-specific Station Corrections (SSSCs) developed for the Preparatory Commission for the Comprehensive Nuclear-Test- Ban Treaty Organization (CTBTO PrepCom) (Firbas *et al.*, 1998), the SALSA3D model with all available data (Ballard *et al.*, 2016a), and the SALSA3D model with no validation events (SALSA3D-NVE). Since the RSTT model is only defined for regional (≤ 15 degrees) distances, it is used only for Pn phases, with *ak135* used for teleseismic P phases. This is how RSTT is designed to be used, and we adhere to that definition.

All relocations are processed using the LocOO3D location algorithm [https://www.sandia.gov/salsa3d/Software.html] (Ballard *et al.*, 2008; Ballard *et al.*, 2009), which permits using combinations of standard 1D lookup tables, the RSTT model, 3D travel-time lookup tables (in GeoTess format (Ballard *et al.*, 2016b)), and full ray bending for the SALSA3D model (Ballard *et al.*, 2009). Depths for all events are fixed during relocation, given the known trade-off between origin time and depth during standard relocation procedures.

To account for *a priori* model errors as defined in each model, we calculate coverage ellipses instead of confidence ellipses. We relocate events, calculating the 95% or 90% coverage ellipses, depending on which models are being compared, to include the use of the *a priori* model errors as defined in each model. The International Data Centre (IDC) for the CTBTO PrepCom runs hypocenter solutions using the 90% coverage ellipse. The SALSA3D models, RSTT model, and CTBTO PrepCom SSSCs include PDU in their relocations, with *ak135* relying on 1DU. Below are the metrics to be considered for evaluating each model's results:

1. Overall location accuracy and precision, where the entire validation set will be measured for median mislocation.
2. Median mislocation relative to the total number of arrivals.
3. Median mislocation relative to maximum azimuthal gap.
4. Pattern of results for "mislocation" grids, or median mislocation for each grid point that is defined as a combination of teleseismic P and regional Pn phases for validation events.
5. Differences of "mislocation" grids between two models. This has the benefit of visually demonstrating reduced mislocation by grid point.
6. One-to-one comparisons of total median mislocation for each event relative to two models.
7. Error ellipse coverage percentage or the percent of the time the GT event falls within the calculated 90/95% coverage ellipse. Shown relative to total number of arrivals and maximum azimuthal gap.

Metric 7 directly tests the effect of the travel time uncertainty (i.e., model error) estimates for each model. For this metric, we are interested in whether the model and its associated uncertainty correctly allow 90/95% of the events to have their error ellipse include the GT location. Similar to Bondár and Storchak (2011), the error ellipse for any relocated event is increased to account for the GT level of the event before determining if the GT event is contained within.

Validation Set 1

In order to validate the SALSA3D model for location of events using P and Pn phases, it is necessary to select validation events that are both globally distributed and that have a sufficient number of phases to be used for location testing. We chose those events that were used for the validation of the Regional Seismic Travel Time (RSTT) model (Myers *et al.*, 2010) in Eurasia as well as in North America (Myers *et al.*, 2011) (Figure 3). The RSTT validation events are identified as ground-truth (GT) events, with epicenter accuracy known to 5 km or better (i.e., GT5) using the Bondár et al. (2004) criteria for estimating potential GT events. By choosing these validation events, we can compare to results when using the RSTT model as well as the 1D *ak135* model (Kennett *et al.*, 1995), both of which have been used by monitoring agencies for routine location of events related to nuclear explosion monitoring.

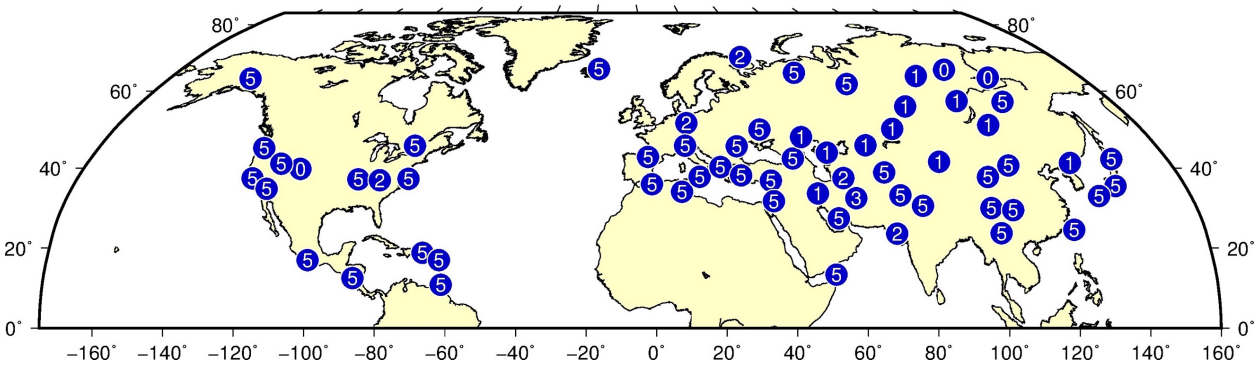


Figure 3. Ground-truth (GT) events chosen for validation of the SALSA3D model. The GT level of each event (in kilometers) is listed on the figure. The events shown were also used for validation of the RSTT model (Myers et al., 2010), providing a consistent comparison with events that have a significant number of regional and teleseismic P arrivals.

The selected validation events were used in the tomographic inversions of SALSA3D as well as RSTT. In order to remove the circularity of using validation events that were also included in the data set for tomography, we perform a new tomographic inversion of SALSA3D, leaving out the selected validation events during the final tomography run (labeled "Relocation Iteration 3" in Ballard et al., 2016b) (denoted SALSA3D-NVE in this article). To demonstrate the effect of leaving out the validation events and compare results to a "circular" test of a model, we will also show validation results when using the full SALSA3D model.

The selected validation events can contain several hundred first-P arrivals, with both regional Pn as well as more standard teleseismic P arrivals. Because of the trend in nuclear explosion monitoring towards regional distance monitoring, one of the questions we wish to address during validation is how well the SALSA3D model handles Pn phases in location, especially when combined with teleseismic phases. The use of regional phases can degrade location accuracy (Myers *et al.*, 2011) when combined with teleseismic phases if the transition from regional to teleseismic is not handled properly in the model. To test the use of regional phases, we develop a "grid" of combined Pn (X-axis) and P arrivals (Y-axis) (Figure 4), with each position in the grid randomly sampling the specified Pn/P proportion from the available phases for a particular event and the shading of the grid point indicating median mislocation. The total number of phases allowed is capped at 20 total phases. From various tests, we chose to randomly sample phases for each grid point and each validation event 50 times. With a total of 63 validation events, a single grid point could contain up to 63 x 50 = 3150 random realizations, producing a "mislocation grid". For a validation test of this type, we wanted grid point populations to be as equal as possible so that the resulting tests will have smoother variations between grid points/total arrivals and are easier to directly compare. However, every validation event did not necessarily have enough total Pn and P phases to successfully sample 50 times per grid point. Hence, we sampled as many times as possible for each Pn/P grid point and validation event (up to the maximum of 50 samples), with the total count of realizations shown in Figure 4 (692,550 total realizations). By randomly sampling different combinations of Pn and P phases (i.e., different combinations of stations), we have the ability to adequately sample events with varying azimuthal gap.

For Validation Set 1, we did not limit stations to any particular network, using any globally-available station. The CTBTO PrepCom SSSCs were only defined for a portion of the stations in the International Monitoring System (IMS) (see Figure 16 of Begnaud et al. (2020b)) and, as such, can't be used with this validation set.

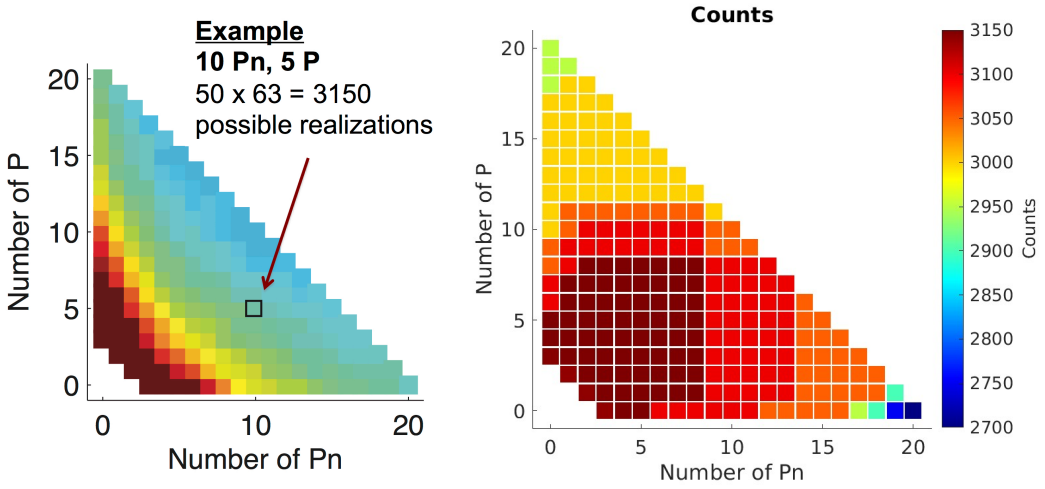


Figure 4. Random realization grid for various combinations of Pn and P arrivals. (left) Example of realizations for a selected Pn/P grid point with colors representing median mislocation, and (right) 2D histogram of counts of numbers of random realizations for each grid point, with a total of 692,550 realizations.

Validation Set 2

Using validation events which have possibly hundreds of available phases suggests these events are larger in magnitude than "typical" events a monitoring agency would encounter during routine operations. In order to test a validation set with more magnitude and arrival variations, we chose to select a second validation set based on GT5 or better events from a specific sparse network, such as the International Monitoring System (IMS).

We identified GT5 or better events using the Los Alamos National Laboratory (LANL) GT database (Begnaud, 2005) which combines seismic bulletins from various sources (e.g., International Data Centre (IDC), International Seismological Centre (ISC), United States Geologic Survey (USGS), other local and regional bulletins) into a reconciled catalog with redundant arrivals removed based on a trusted hierarchy of sources. The events in the LANL reconciled database are tested against GT criteria (e.g., Bondár *et al.*, 2004; Bondár and McLaughlin, 2009, etc.) or identified from other information (i.e., known source locations). Those events that are GT5 or better are then identified that also have IMS station arrivals from the CTBTO PrepCom Reviewed Event Bulletin (REB) produced by the IDC. These GT5 or better events that have a minimum of three P and/or Pn IMS arrivals are collected for the new validation data set.

The second validation data set contains 2302 GT5 or better events with epicenters determined from the LANL GT database and arrivals pulled only from IMS arrivals (Figure 5). This collection has the benefit of sampling more of a realistic range of magnitudes and hence is more representative of "real" events, as compared with a sub-sampled set of events of larger

magnitude with atypically high signal-to-noise ratio waveforms used for arrival picking. Due to the strict definition of GT5 from Bondar et al. (2004) (phases from a local network, detected within 250 km and 10 km), most of the events in this second validation set are concentrated in Europe, Japan, and North America, where there are dense networks of non-IMS stations defining the GT5 or better events. The nature of these REB GT5 or better events is that they are also included in the SALSA3D model. However, these GT events were part of those left out of the "pdu2020012DU" version of the RSTT model and held for validation use (Begnaud *et al.*, 2020a; Begnaud *et al.*, 2020b). Thus, the validation tests will be circular for SALSA3D, but a mostly consistent comparison can be made. It would not be valuable to leave out this number of events from the tomographic inversion and still being able to test the variations in magnitude and/or number of arrivals.

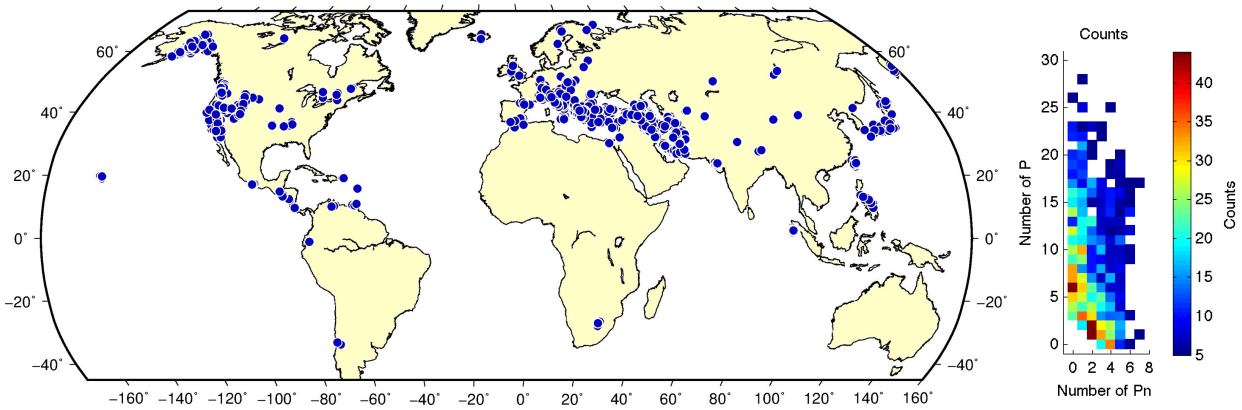


Figure 5. (left) Second validation set selected from GT5 or better events from LANL GT database that have CTBTO REB Pn/P arrivals (2302 events). (right) 2D histogram of counts of combinations of Pn/P arrivals for validation events. Note the variable counts of these events, with the highest numbers of events with smaller numbers of total arrivals.

As stated in the description of Validation Set 1, we are using the relatively static IMS network for model validation and comparison and can use the CTBTO PrepCom's SSSCs. Other 3D models could be used directly in our locator to compare to SALSA3D if both the model and ray-tracing algorithm were available and could be integrated into the locator. For other 3D models to be used for comparison, we could rely on a travel-time ray-tracing code developed for such models to pull out predictions and use GeoTess (Ballard *et al.*, 2016b) to create 3D travel-time lookup surfaces for IMS to compare relocation results. Rowe et al. (2009) clearly showed the widely known result of trying to compare travel time calculations and tomography models using a ray tracing that was not used for that model's development. Travel time calculations MUST use the same ray tracing algorithm between the tomography model and the event locator or results will not be consistent. By using a ray-tracing algorithm for calculating travel times that is distributed with a given 3D model and populating 3D travel-time lookup tables, we could fairly compare the relocation results to other 3D models within the same event location algorithm.

To fairly compare SALSA3D with CTBTO PrepCom SSSCs, we also created 3D travel-time lookup surfaces for SALSA3D and the IMS network, using a tessellated grid spacing of 1 degree. Since we have PDU values available for SALSA3D, we created lookup surfaces to include those

uncertainty values as well as another set of SALSA3D surfaces with 1DU values (labeled SALSA3D-1DU) to demonstrate the effect of the PDU.

Validation Results

Validation Set 1

Figure 6 shows the median mislocation and event count relative to the total number of Pn/P arrivals for each of the tested models/versions in Validation Set 1. Also labeled are the overall median mislocation values for each model. The full SALSA3D model shows a significant improvement over *ak135* (by 29.6%) and the RSTT (14.5%) models, for every number of arrivals. The percent improvement for SALSA3D (full) is greatest for when there are only three total arrivals (ak135 – 34.6%, RSTT – 21.5%) and smallest for 20 total arrivals (ak135 – 29.3%, RSTT – 13.7%), as expected. For all numbers of arrivals, the SALSA3D-NVE version of the 3D model shows virtually identical results when compared to the full SALSA3D model.

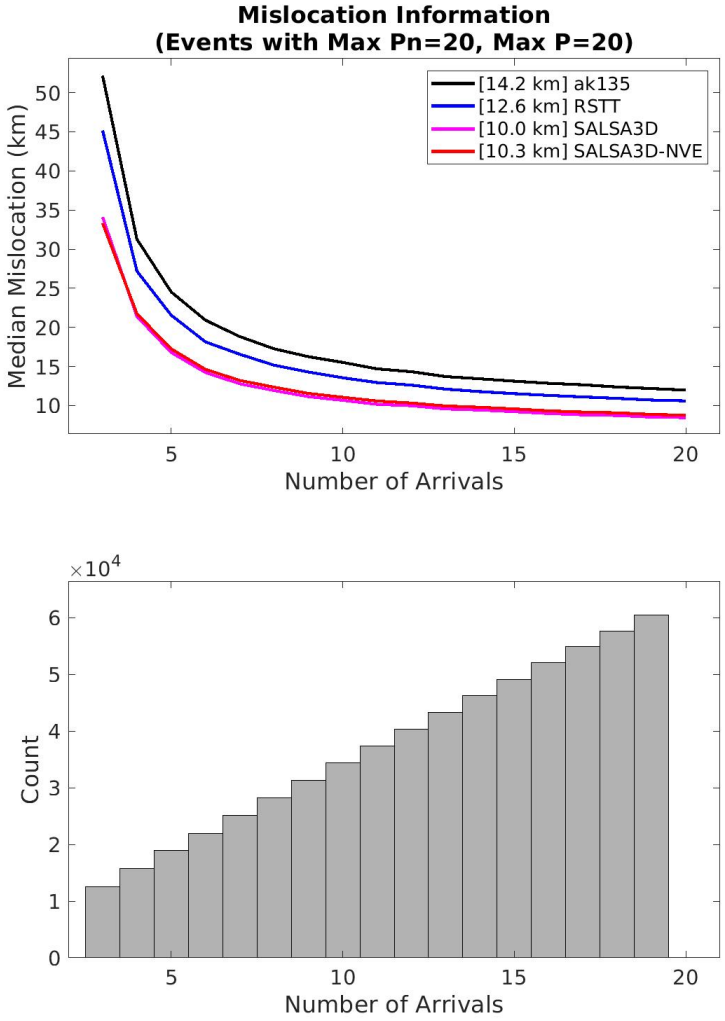


Figure 6. (top) Median mislocations for all validation events and all random realizations, relative to total number of P/Pn phases (all possible P and Pn phases). (black) ak135, (blue) RSTT, (magenta) SALSA3D (all events used during tomography), (red) SALSA3D-NVE (no validation events using during

tomography). Overall median mislocation is labeled for the entire validation set. Note the results for SALSA3D and the SALSA3D-NVE (no validation events) model versions are practically identical, with the SALSA3D-NVE version have a slightly higher median mislocation. (bottom) Histogram of the number of events with the specified arrivals.

To best view the relocation results for each model and all realizations of P and Pn phases, we created "mislocation grids" (Figure 7). These mislocation grids allow us to better understand the behavior of each model as the number and types of phases vary. For realizations of less than about 5-6 total phases, mislocation is high, typically greater than 20 km. The use of the SALSA3D model clearly results in reduced mislocation values when compared to ak135 and RSTT in this 5-6 total phase range. In addition, the SALSA3D model shows the lowest mislocation results concentrated at approximately 13 Pn with 8 P phases (slightly favoring Pn phases) while RSTT shows the most favorable mislocations near the maximum number of allowed Pn phases. This is not surprising given that the RSTT model is designed specifically for regional phases. The ak135 model shows minimum mislocation values at roughly equal numbers of P and Pn phases.

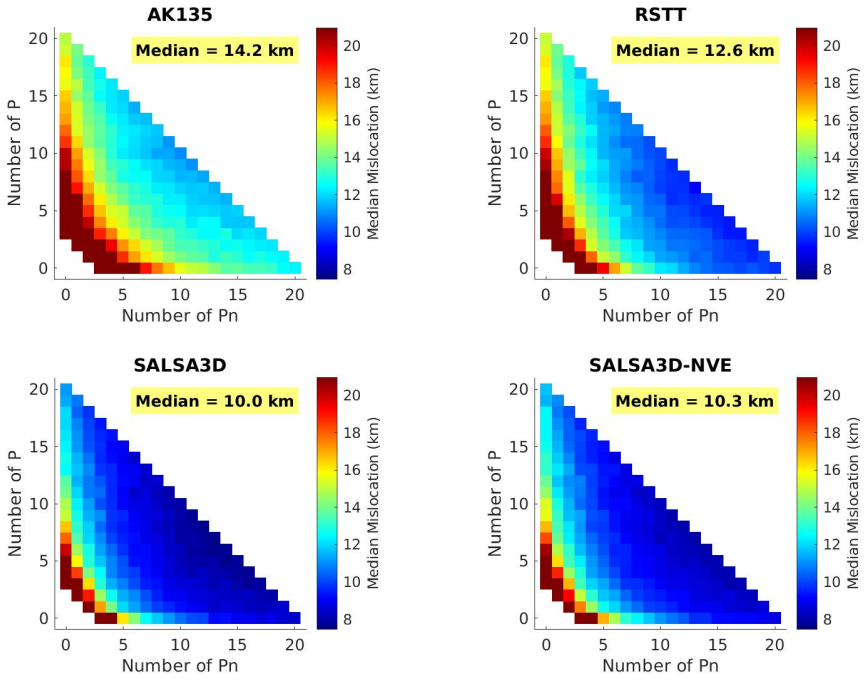


Figure 7: Median mislocations, relative to the ground-truth epicenter, for each tested model and each number of P and Pn arrivals, producing "mislocation grids" for a minimum of 3 total phases to a maximum of 20 total phases. For each model, the overall median mislocation is shown. For realizations of less than about 5-6 phases, mislocation is high, typically greater than 20 km. The use of the SALSA3D model (NVE denotes "no validation events") clearly results in reduced mislocation values when compared to ak135 and RSTT in this 5-6 total phase range. In addition, the SALSA3D model shows the lowest mislocation results concentrated at approximately 13 Pn with 8 P phases, while RSTT shows the most favorable mislocations near the maximum number of allowed Pn phases. The ak135 model shows minimum mislocation values at roughly equal numbers of P and Pn phases.

Figure 8 show the difference between the mislocation grids for each model (ak135, RSTT, SALSA3D) compared to every other model. This allows for better direct model comparison, especially for certain phase combinations. Overall, RSTT has lower median mislocation values for each phase combination (average 1.8 km) as compared to ak135. The exception is when there are no Pn phases; here RSTT essentially becomes ak135, as RSTT is not designed for P phases (i.e., teleseismic). SALSA3D shows a significant reduction in mislocation over ak135 (average 4.7 km) when total phases are less than approximately 10-15, while still showing reduced mislocation for the entire grid. SALSA3D is also showing reduced mislocation (average 3.0 km) over RSTT, with significant reduction in mislocation in the grid near a majority of P phases. SALSA3D shows reduced mislocation over RSTT in every grid cell except the 3-Pn phase cell, where RSTT has a lower mislocation.

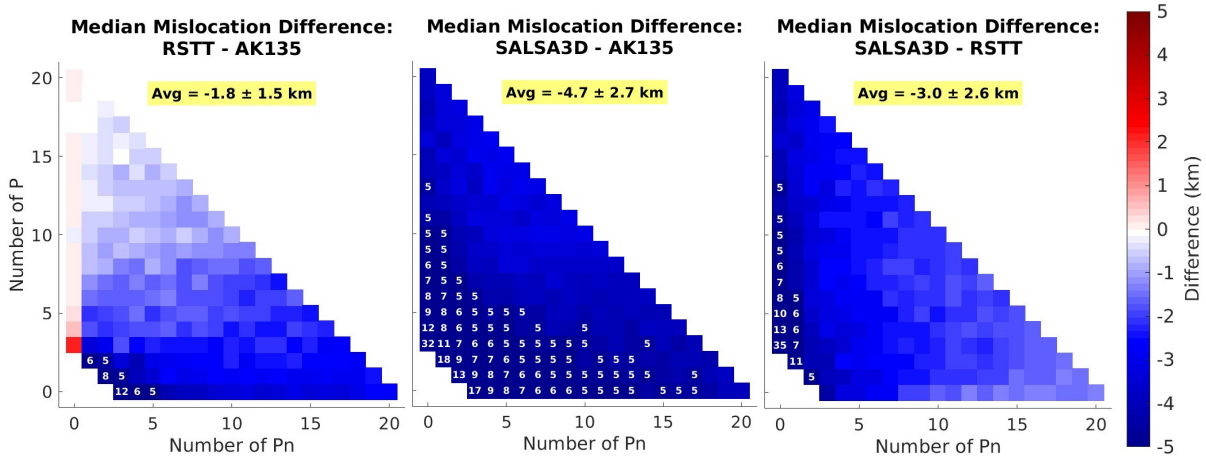


Figure 8: Difference in mislocation grid values for each model combination (not including the SALSA3D-NVE). (left) RSTT-ak135, (middle) SALSA3D-ak135, (right) SALSA3D-RSTT. Blue colors indicate the first model has lower median mislocation values than the second. Median mislocation differences of ≥ 5 km are indicated with the difference value shown for each grid cell. SALSA3D is showing the greatest reduction in mislocation over ak135 and RSTT, with the difference with ak135 being greatest and significant when total phases are less than about 10-15.

Figure 9 demonstrates the effect on relocations of the tomography model with no validation events (to remove validation circularity) compared to including all validation events. Leaving out the validation events during tomography does result in an actual effect on the relocations with a 0.3 km increase in overall mislocation value. The 0.3 km increase is consistent almost throughout each realization combination of P and Pn phases. The no-validation-events model (SALSA3D-NVE) shows reduced mislocation only when total phases are 3-4, with the most significant reduction at 0 Pn and 3 P phases.

Station azimuthal gap has one of the most significant effects on event mislocation compared to many other effects (e.g., Bondár *et al.*, 2004), including the individual model used. No model, no matter how accurate, will entirely remove the effect of azimuthal gap, although it can diminish it to some degree. Figure 10 shows median mislocation values for each tested model and all phase realizations, relative to azimuthal gap. Bin size is 20°, with an overlap of 10°. The accuracy improvement of each model is most obvious for the largest azimuthal gaps (> 250°),

with SALSA3D having the lowest mislocation. The average azimuthal gap is ~ 143° at which the median mislocation is 14.3 km for ak135, 12.5 km for RSTT, 9.7 for SALSA3D, and 10.1 km for SALS3D-NVE. At a less common nor ideal azimuthal gap of 200° (i.e., mean + 1-sigma), the median mislocations for ak135, RSTT, SALSA3D, and SALSA3D-NVE are 19.4, 18.1, 14.4, and 15.1 km, respectively.

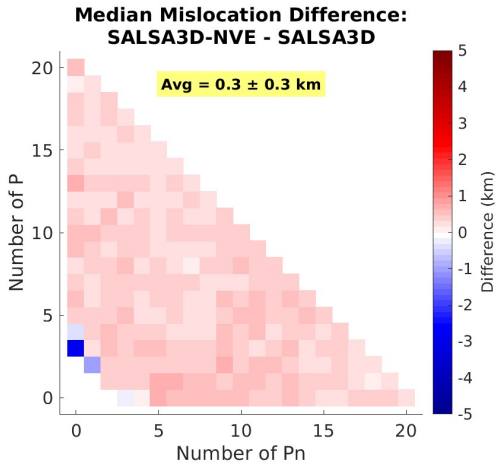


Figure 9: Mislocation difference of the SALSA3D-NVE (no validation events) with the SALSA3D model. When validation events are left out of tomography to remove validation circularity, the relocation results show a consistent increase in mislocation of 0.3 km, with that value being consistent in almost all realization combinations of P and Pn phases.

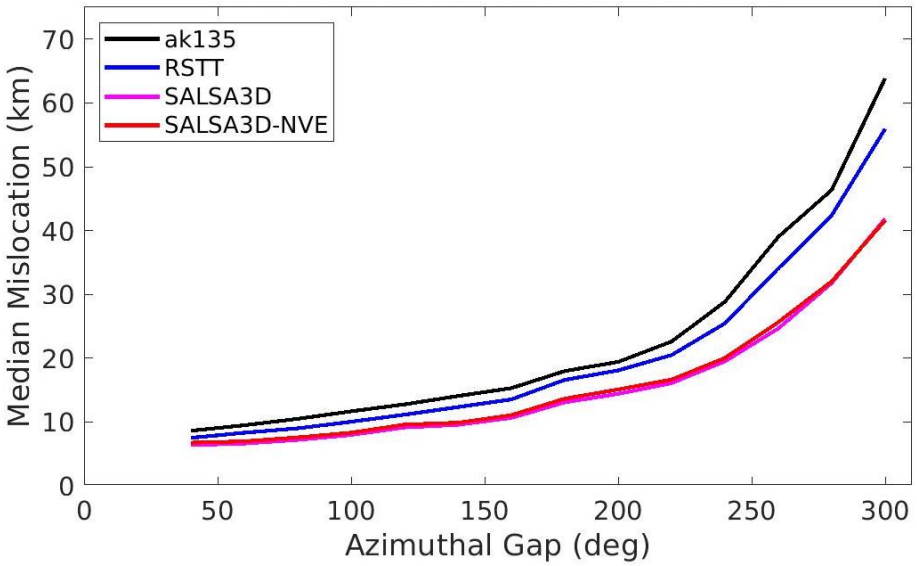


Figure 10: Median mislocation for each model and all phase realizations, relative to azimuthal gap. The results are binned by 20° of azimuthal gap, with an overlap of 10°. Refer to Figure 6 for color model assignments.

Figure 11 shows one-to-one plots of models and events, comparing the median mislocations for each event separately along with the median absolute deviation from the median. On the plots, we show the number of events where a specific model shows lower median mislocations. The

SALSA3D model shows a greater number of events where median mislocations are lower than the model being compared against. Compared to ak135, SALSA3D had 55 (SALSA3D-NVE) and 58 events of the 63 total events with lower median mislocations. Compared to RSTT, SALSA3D had 53 (SALSA3D-NVE) and 54 events with lower median mislocations. Comparing the two versions of SALSA3D, the model with all events in the tomography (SALSA3D) edges out the model where the validation events were removed (SALSA3D-NVE).

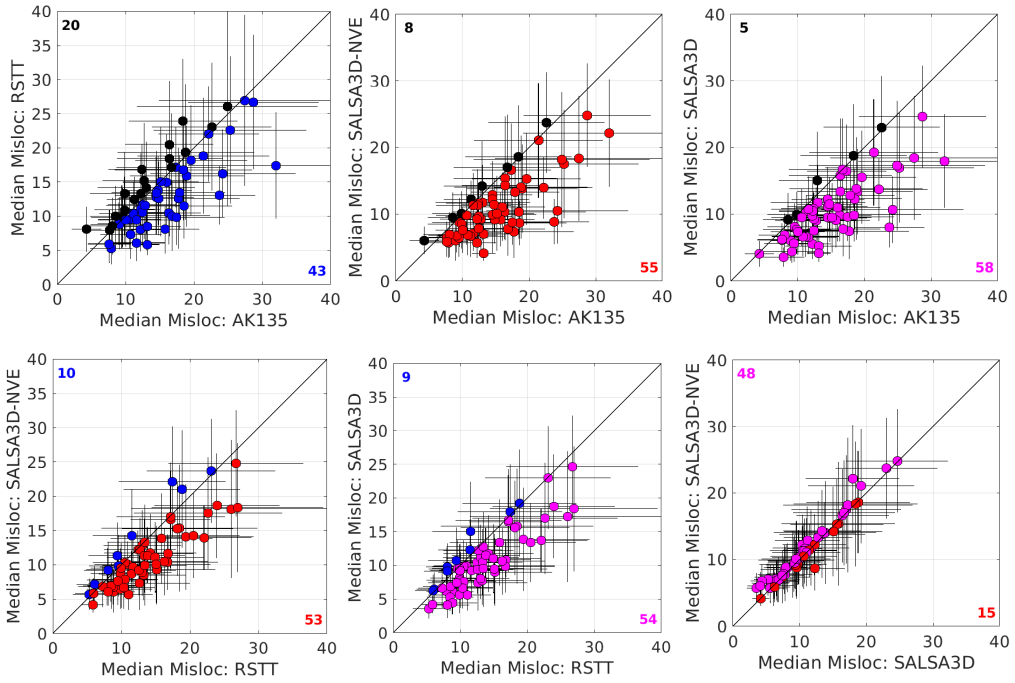


Figure 11: One-to-one plots of median mislocation for each of the 63 validation events, comparing median values and median absolute deviation from the median (MAD1) for two models at a time. Refer to Figure 6 for color representations. Values in opposite corners of each square indicate the number of events where a specific model showed lower median mislocation. The SALSA3D model shows a greater number of events where median mislocations are lower than the model being compared. 55 (SALSA3D-NVE) and 58 (SALSA3D) as compared to ak135. 53 (SALSA3D-NVE) and 54 (SALSA3D) compared to RSTT. Comparing the two versions of SALSA3D, the model with all events in the tomography edges out the model where the validation events were removed.

One of the primary goals of developing the SALSA3D global velocity model was to fully solve for the 3D covariance matrix in order to be able to calculate travel-time path uncertainty throughout the globe (Ballard *et al.*, 2016a). Having more realistic and proper estimates of travel-time uncertainty can affect relocation results as phases will be weighted based on PDU values rather than simple 1DU (e.g., ak135). For events with small azimuthal gaps, using PDU may not greatly affect the determination of the epicenter. However, as azimuthal gap increases, mislocation increases, and variable phase weighting from PDU could be a controlling factor on the resulting relocation.

The metric for coverage ellipse percentage is the percent of the time the GT location is within the coverage ellipse. For a 95% coverage ellipse, the GT location should fall within the error ellipse 95% of the time. Figure 12 shows the ellipse percentage for the tested models relative to

the number of arrivals. For the realizations in Validation Set 1, the overall ellipse percentages are very similar, at ~94%. The SALSA3D model shows a slight improvement over the ak135 model, with the RSTT model consistently overshooting the 95% goal with a median value of 97.2%. SALSA3D is demonstrating a flatter, or more consistent, curve at higher number of arrivals. All models exhibit a reduction in percentage at the inflection point of the curve, near 6-7 arrivals.

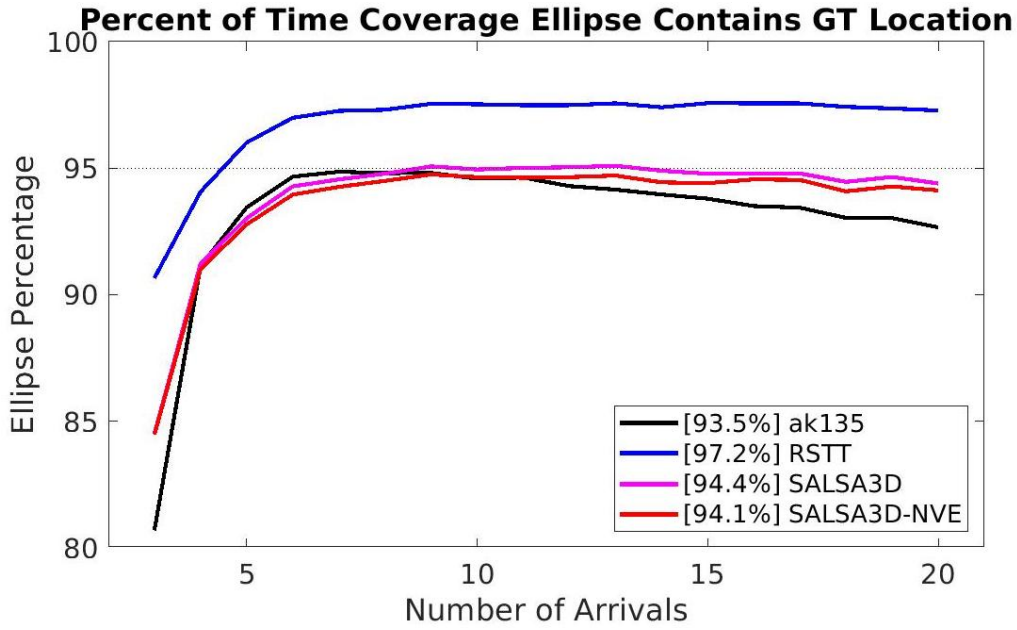


Figure 12: Percent of time the 95% coverage ellipse contains the GT location relative to number of phases, for all possible phase realizations. Refer to Figure 6 for model color representations. The SALSA3D model has a flatter or more consistent curve with increasing numbers of arrivals. However, at the inflection point at around 6 arrivals, all models exhibit a reduction in the ellipse percentage. Overall, the percentage for most models are around 94%, with SALSA3D showing a minor improvement over ak135. However, the RSTT model consistently overshoots the 95% goal for Validation Set 1.

The total number of arrivals does not always correspond to the azimuthal gap for an event. Figure 13 shows the ellipse percentage relative to azimuthal gap, using a 10° azimuthal gap bin size and bin overlap of 5°. At smaller azimuthal gaps where there is good station coverage, SALSA3D tends to be greater than the 95% target with ak135 below slightly below. RSTT has PDU like SALSA3D but is again showing a consistently high value, overshooting the 95% goal. The majority of events have azimuthal gaps from 70°-150°. As azimuthal gap increases, all models demonstrate a consistent reduction in ellipse percentage. At ~280°, all models fall-off in ellipse percentage, with SALSA3D and RSTT showing a rebound at ~340°.

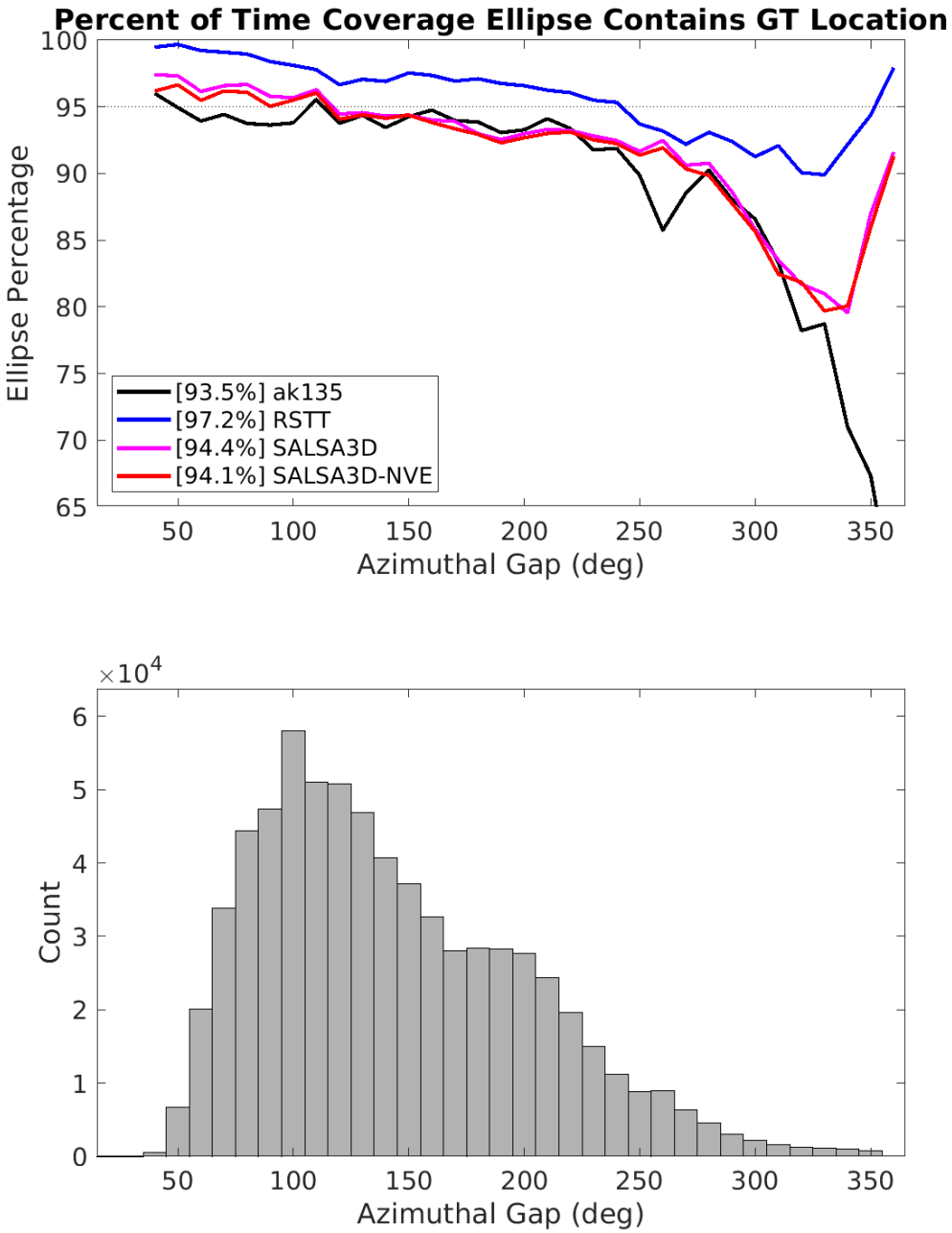


Figure 13: Percent of time the 95% coverage ellipse contains the GT location relative to azimuthal gap, for all possible phase realizations. Refer to Figure 6 for model color representations. All models exhibit a consistent download slope with increasing azimuthal gap. The SALSA3D models (PDU and 1DU) are near the 95% goal, but the RSTT model appears to overshoot to 97.2%.

Validation Set 2

The median mislocation results for Validation Set 2 are shown relative to the total number of P and Pn arrivals in Figure 14. Mislocation results are similar for ak135 (16.1 km), CTBTO PrepCom SSSCs (15.5 km), and RSTT (15.2 km). The Pn phase is not as prevalent for Validation Set 2 as for Validation Set 1, owing to the more limited availability of that phase (max of 7) for the IMS network (Figure 5) and the selected validation events. SALSA3D mislocation results show

significantly reduced values when using PDU (11.3 km) and 1DU (10.5 km). Above ~25 arrivals, SALSA3D is similar to the other compared models, but still showing reduced mislocation values overall. Figure 15 shows mislocation results relative to azimuthal gap. Again, SALSA3D is demonstrating consistently reduced mislocation results over most of the azimuthal gap range.

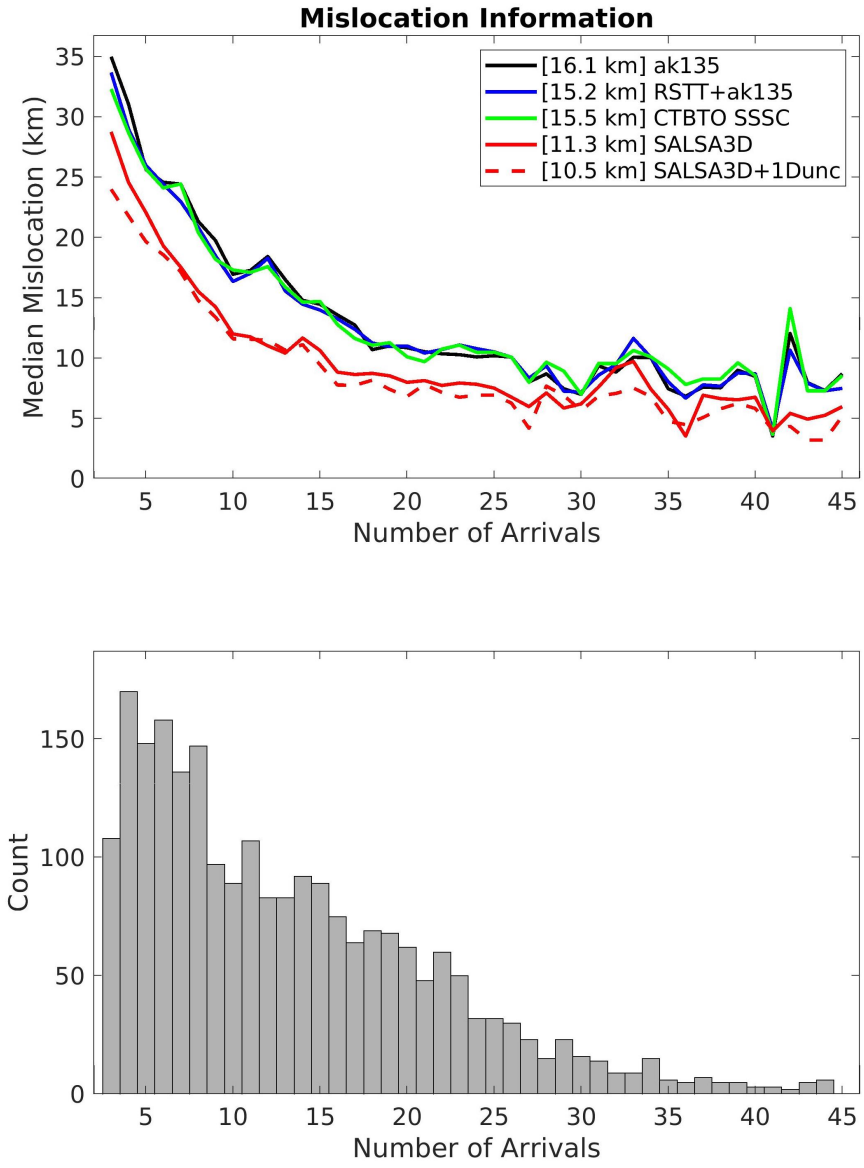


Figure 14: (top) Median mislocations relative to total number of P and Pn arrivals for Validation Set 2 (overall values for each model shown). The ak135 (black), RSTT+ak135 (blue), and CTBTO PrepCom SSSCs (green) display similar and consistent mislocation results over the range of available arrivals. SALSA3D results for path-dependent uncertainty (solid red) or 1D distance-dependent uncertainty (dashed red) are significantly lower than the other models. (bottom) Histogram of the number events relative having a certain number of arrivals.

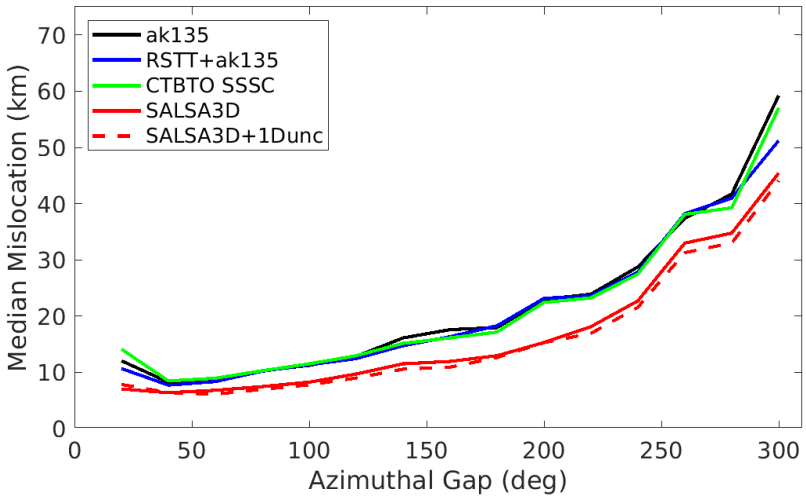


Figure 15: Median mislocation results relative to azimuthal gap. As for Figure 14, the SALSA3D model has significantly reduced values compared to the other models shown, with consistent differences for most of the range of azimuthal gap.

Unlike for Validation Set 1, mislocation grids for Validation Set 2 are limited to what P and Pn phases are available for the selected events, with the Pn phase limited to a maximum of 7, with the majority of maximum values around 5 (Figure 16). For all tested models, median mislocation values are higher when there are fewer numbers of total arrivals ($\lesssim$ 6 arrivals). The SALSA3D model shows reduced mislocation values compared to other models particularly when there are fewer than ~7 total arrivals but a majority of Pn arrivals (lower right of each grid). SALSA3D is showing reduced mislocation when events have a larger majority of teleseismic P arrivals.

In order to compare models in detail, we difference the grids to highlight mislocation for specific phase combinations (Figure 17). Results for SALSA3D demonstrate significant reduction in mislocation values, many above the 5 km difference labeled in the figure, for ak135, RSTT, and CTBTO PrepCom SSSCs.

Figure 18 shows one-to-one plots of models and events for Validation Set 2, comparing the median mislocations for each event separately. On the plots, we show the number of events where a specific model shows lower median mislocations. The SALSA3D model shows a greater number of events where median mislocations are lower than the model being compared. When comparing SALSA3D to SALSA3D-1DU, the version using the 1DU has the advantage with 1332 events over SALSA3D with 970.

We validate again for the percent of the time the GT event falls within the error ellipse of the relocated event. For Validation Set 2, we chose to relocate and solve using a 90% coverage ellipse as this is the operational behavior for the IDC of the CTBTO PrepCom. Figure 19 shows the error ellipse percent relative to the total number of arrivals. Overall, all models demonstrate somewhat consistent values across the total number of arrivals. SALSA3D and RSTT+ak135 are the closest to the 90% goal (87.8% and 88.5%, respectively), as they both use more appropriate PDU values. The CTBTO PrepCom SSSCs (82.8%) do have PDU values as well,

but are not calculated for all IMS stations (since many IMS stations were certified after creation of the SSSCs), defaulting to ak135 1DU, and shows overall values similar to ak135 (82.9%). The version of the SALSA3D model with only 1DU has much lower percentage values (68.4%) suggesting the 1DU is not accounting for the required travel-time uncertainty to meet the 90% goal.

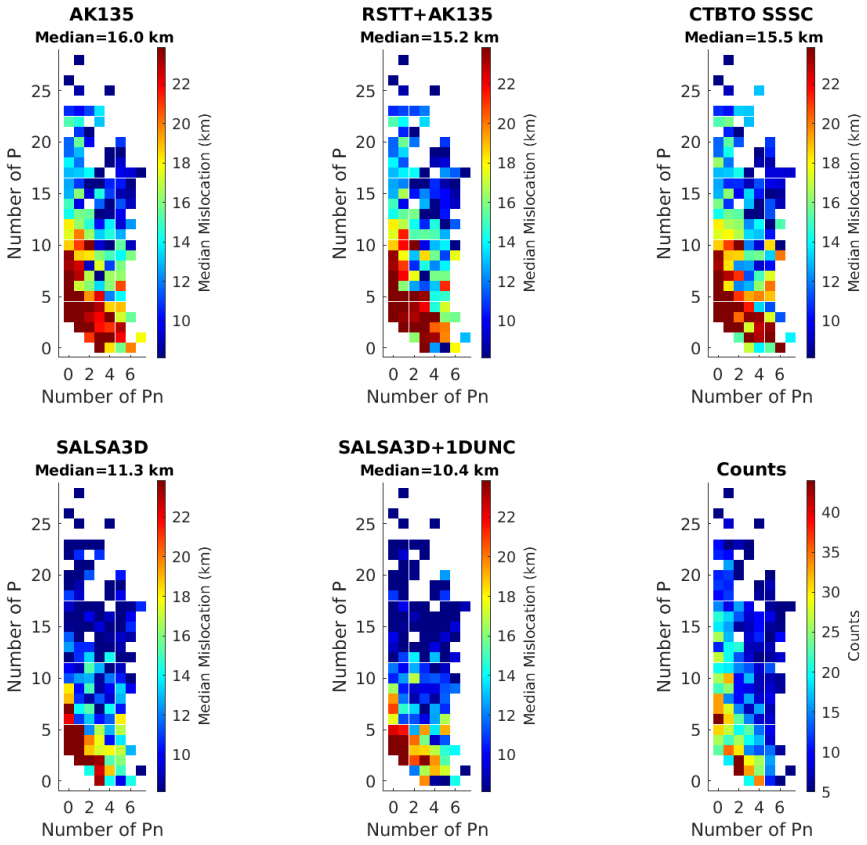
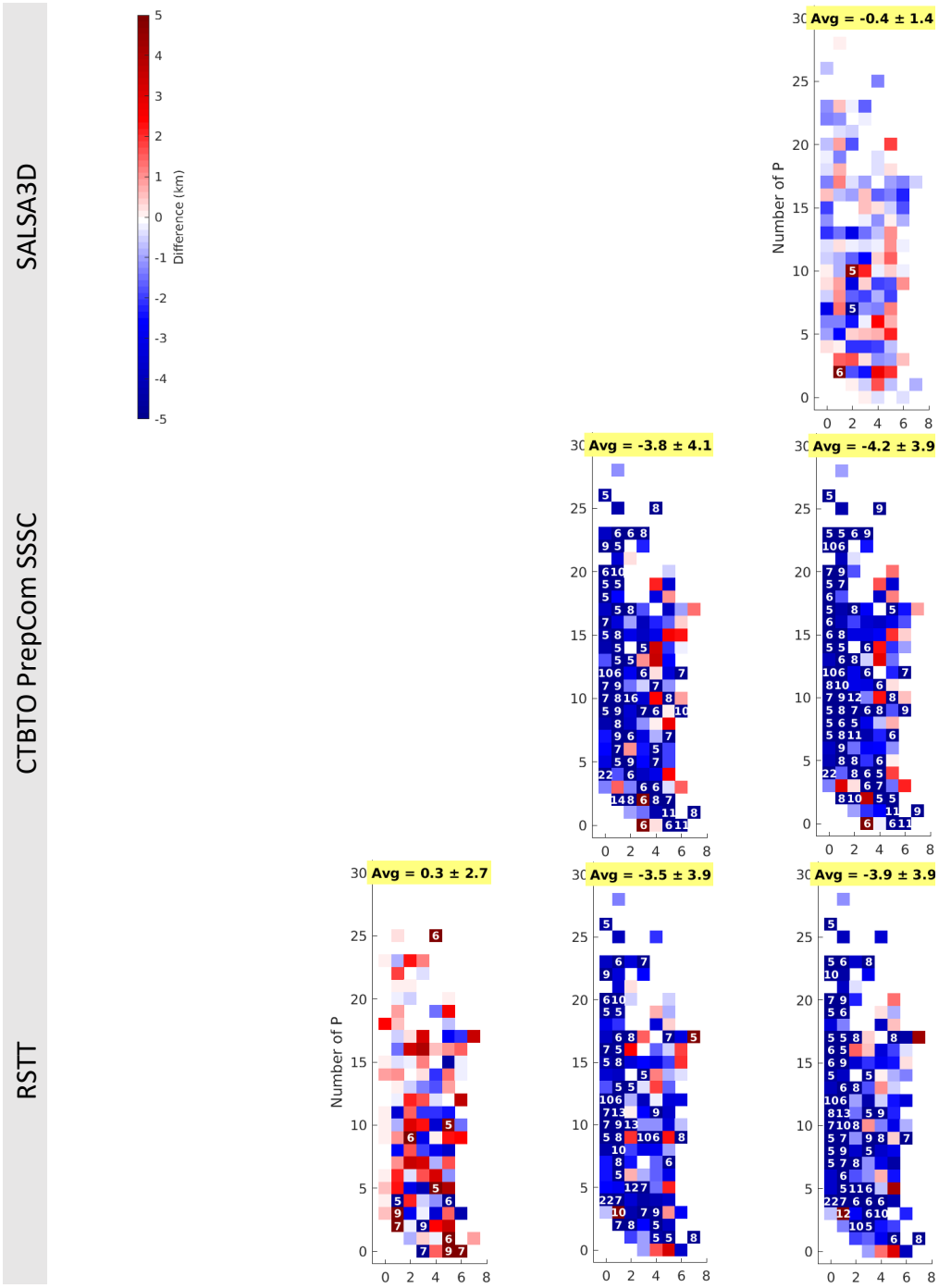


Figure 16: Mislocation grids for Validation Set 2 and the tested models shown, with number of P on the vertical axis and number of Pn on the horizontal axis. Total number of events for each combination of phases is shown as "Counts" in the lower right grid. Colors represent the median mislocation for all events with phase combinations defined at that grid point. The colors for mislocation grids represent median mislocation (km). All tested models exhibit larger mislocations values for fewer phases. SALSA3D models show reduced mislocation for greater number of P/Pn phases.

Figure 20 shows the error ellipse percent relative to the event azimuthal gap for Validation Set 2. The error ellipse percent for this validation set results in more expected patterns compared to Validation Set 1 based on use of PDU (e.g., SALSA3D, RSTT, CTBTO PrepCom SSSCs) over 1DU (e.g., ak135, SALSA3D+1DU). The models using some form of PDU result in ellipse percentages much closer to the 90% goal than those using 1DU. The values for ak135 (82.9%) though are much closer to the goal than SALSA3D-1DU. SALSA3D (87.8%) and RSTT+ak135 (88.5%) come the closest to the 90% goal. As ak135 does approach the percentage goal, it is aiding the RSTT+ak135 model for teleseismic P phases (RSTT is not defined for teleseismic phases), resulting in the near 90% value.


SALSA3D
CTBTO PrepCom SSSC
RSTT
Difference (km)
Avg = -0.4 ± 1.4
Avg = -3.8 ± 4.1
Avg = -4.2 ± 3.9
Avg = 0.3 ± 2.7
Avg = -3.5 ± 3.9
Avg = -3.9 ± 3.9
Number of P

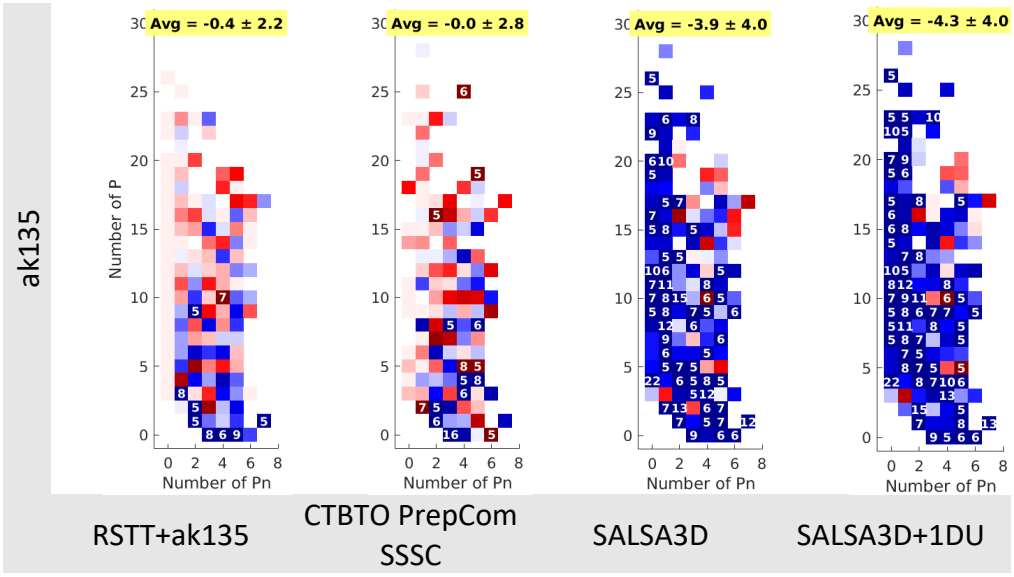


Figure 17: Differences between model mislocation grids showing COLUMN – ROW. Blue colors indicate model in COLUMN has lower mislocation than column in ROW. Overall mean and standard deviation of the mislocation differences for each grid point. If absolute difference is ≥ 5 km, the difference value is shown in the grid. The SALSA3D model (both path-dependent (PDU) and 1D distance-dependent (1DU) uncertainty) display significantly lower mislocation values, especially for events with a majority of teleseismic P phases. SALSA3D+1DU is demonstrating a reduced overall mislocation value of ~0.4 km over SALSA3D with PDU.

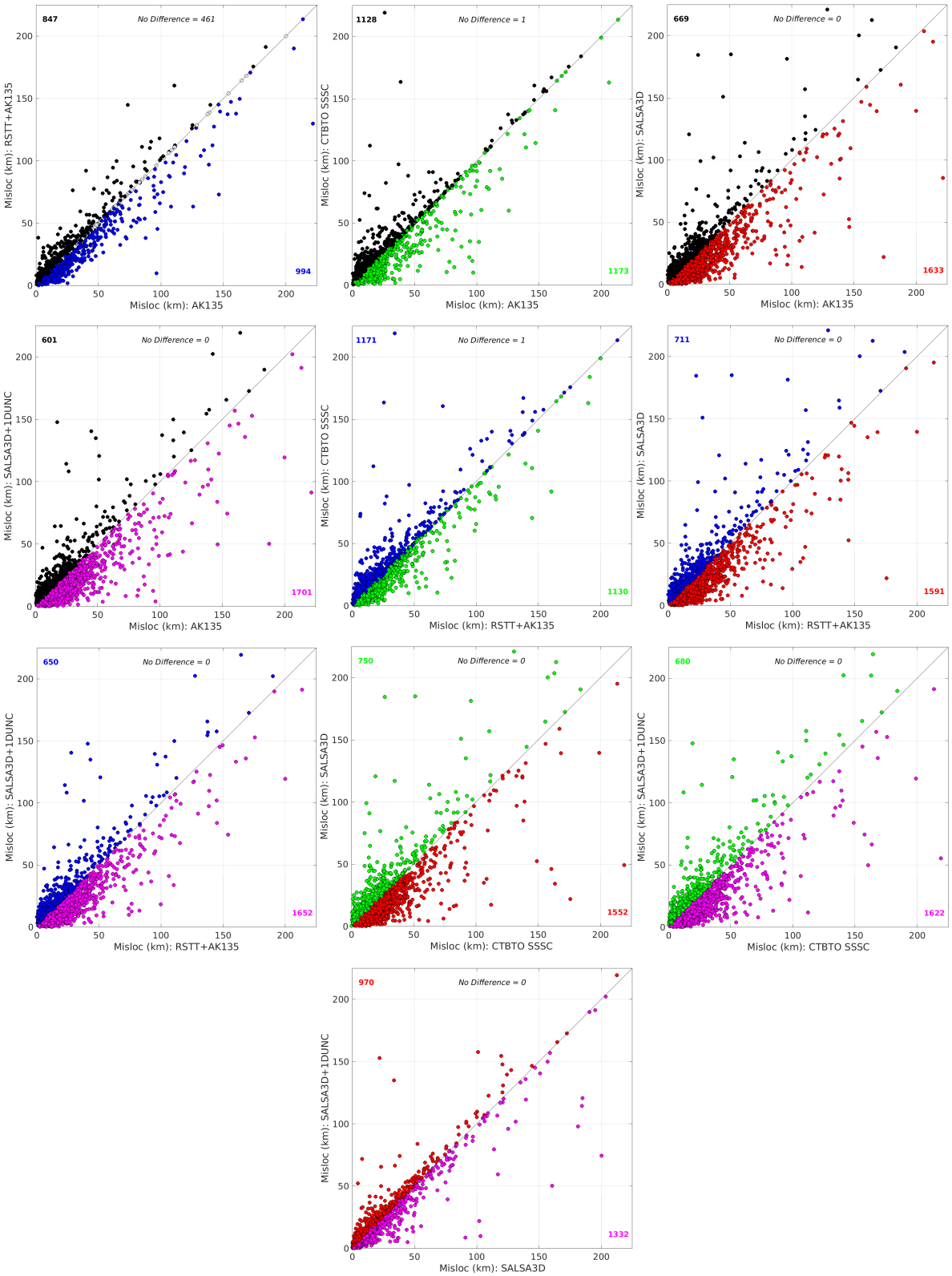


Figure 18: One-to-one event mislocation plots for the tested models: (black) ak135, (blue) RSTT+ak135, (green) CTBTO PrepCom SSSCs, (red) SALSA3D, (magenta) SALSA3D+1DU. The number of the validation

events using a particular model and resulting in a lower mislocation compared to the second model is shown in the corners with the corresponding color listed above. The exact one-to-one line is shown as well as those events where the mislocation was the same for the compared models.

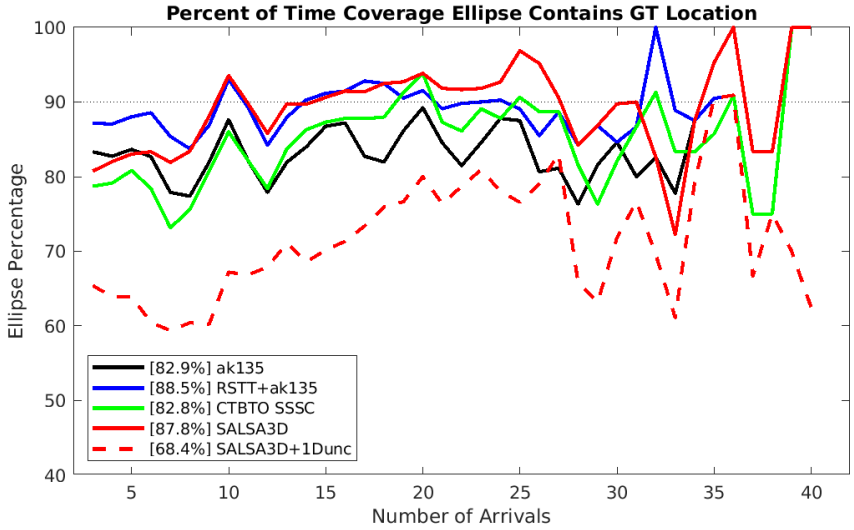


Figure 19: Percent of the time the 90% coverage ellipse contains the GT location for the selected models and total number of arrivals in Validation Set 2. See Figure 14 and/or the legend in the top figure for colors. The 90% line is shown as the percent goal for each model. Overall, all models demonstrate somewhat consistent values across the total number of arrivals. SALSA3D and RSTT+ak135 are the closest to the 90% goal, as they both use appropriate path-dependent uncertainty values. The CTBTO PrepCom SSSCs do have path-dependent uncertainties as well, but are not calculated for all IMS stations, defaulting to ak135 1D distance-dependent uncertainty (1DU), and shows overall values similar to ak135. The SALSA3D-1DU has much lower percentage values suggesting the 1DU is not accounting for the required travel-time uncertainty to meet the 90% goal.

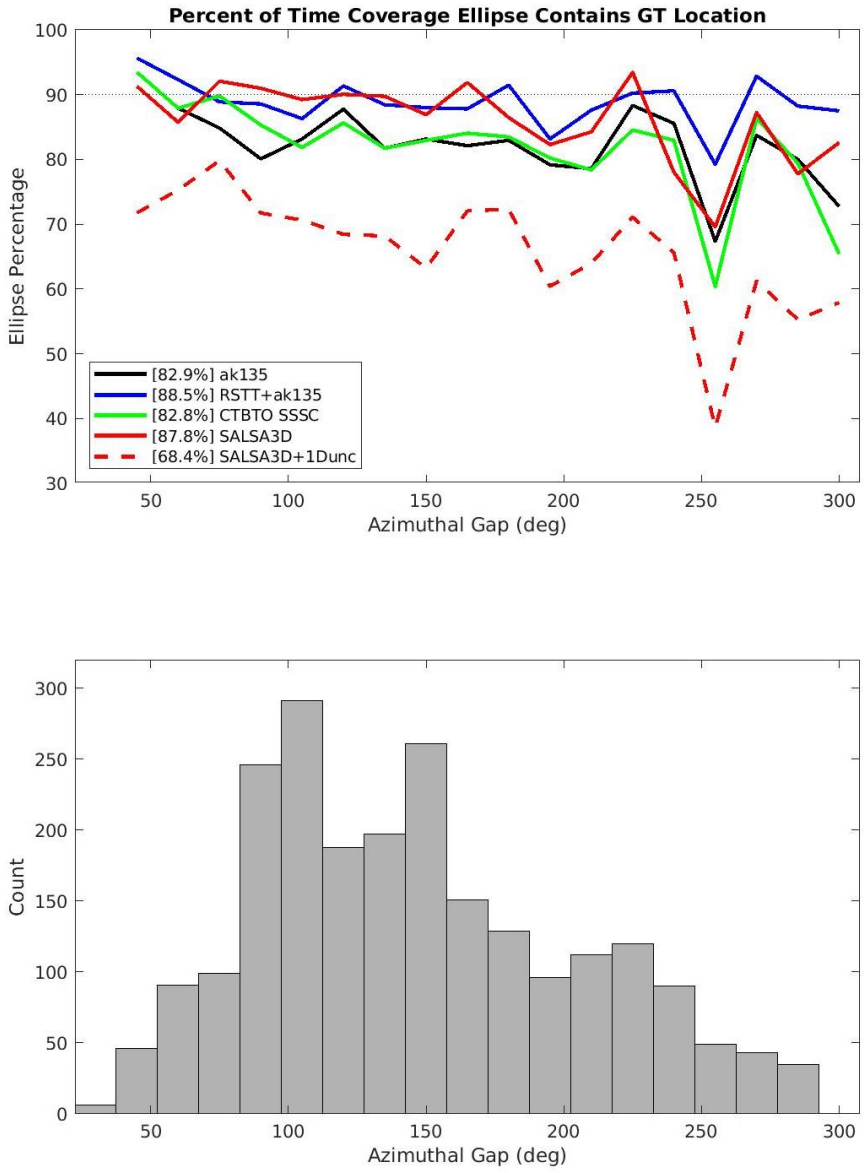


Figure 20: (top) Percent of the time the 90% coverage ellipse contains the GT location for the selected models and azimuthal gap in Validation Set 2. Bin size is 15 degrees. See Figure 14 and/or the legend in the top figure for colors. The 90% line is shown as the percent goal for each model. (bottom) Histogram of the number of events within each azimuthal gap bin. The error ellipse percent for this validation set results in more expected patterns based on use of path-dependent uncertainty (PDU) (e.g., SALSA3D, RSTT, CTBTO PrepCom SSSCs) over 1D distance-dependent uncertainty (e.g., ak135, SALSA3D+1DU). The models using some form of PDU result in ellipse percentages much closer to the 90% goal than those using 1DU. (bottom) Histogram of the number of events for the given azimuthal gap.

Discussion and Conclusions

The development and validation of the SALSA3D global model has shown significant reduction in the overall mislocation of events over other 1D (ak135) and higher-dimensional (2D, 3D) (RSTT, CTBTO PrepCom SSSCs) models. Even when comparing to a version of SALSA3D with no validation events (SALSA3D-NVE), the difference compared to the full data version is minor, a difference of only 0.3 km overall. For Validation Set 1, SALSA3D results in lower mislocation

values, particularly for realizations with a majority of teleseismic P phases. This is not surprising given that RSTT uses ak135 for teleseismic P phases. However, SALSA3D also shows reduced mislocation values for realizations with predominantly regional Pn phases.

For Validation Set 2, which contains “actual” events and their associated P/Pn phases (as opposed to the multiple random realizations in Validation Set 1), SALSA3D also results in significantly reduced median mislocation values over the compared models, particularly for those events using more teleseismic P phases. Note that many events with a predominance of regional Pn phases also show SALSA3D with the advantage, but to a lesser extent.

One of the recent primary research goals for 3D model development has been the calculation of some form of 3D PDU in order to better estimate the correct travel-time uncertainty for any given path in a model. However, for SALSA3D the use of PDU compared to 1DU results in a slight increase in median mislocation. We speculate that this may be because including PDU with the 3D model will significantly change the arrival weights (inversely proportional to uncertainty) used during relocation of a given event. 1DU travel-time uncertainty values are the same for any station at the same distance from an event, but this will not be the case for PDU uncertainties. Perhaps the PDU is adding a less smooth or consistent set of model errors values, resulting in the observed mislocation increase.

For the thousands of random realizations in Validation Set 1, the use of 1DU versus 3D PDU does not appear to be a significant controlling factor for the percent of time the GT location is contained within the 95% coverage ellipse. However, Validation Set 2 demonstrates more of the expected results when using PDU versus 1DU. Validation Set 1 is actually only 63 events with selections of various associated teleseismic P and regional Pn phases, resulting in 692,550 total realizations. All the realizations of the validation events are essentially tracing the same paths and appear to result in more consistent ellipse percent values than for Validation Set 2, which has 2302 distinct events with varying numbers of P and Pn arrivals.

The error ellipse percent plot relative to azimuthal gap for Validation Set 1 in Figure 13 clearly shows a consistent drop in ellipse percent with increasing azimuthal gap (usually associated with fewer phases (Figure 12)). Only the SALSA3D model versions and RSTT+ak135 demonstrate a sharp recovery for the last azimuthal gap bin. This drop in ellipse percentage is not readily evident for Validation Set 2 (Figure 20), except for SALSA3D-1DU. Results for Validation Set 2 include less data and exhibit more fluctuation over azimuthal gap than for results using Validation Set 1. The use of PDU in SALSA3D and RSTT+ak135 appear to mostly remove this ellipse percent slope, but not completely. The simple use of a linear assumption in calculating an error ellipse and/or the 2D ellipse representation of the epicenter error may not be adequate to completely account for poor azimuthal gap coverage. While 2D and 3D models with PDU help to correct this pattern, the remaining ellipse percent slope with increasing azimuthal gap or decreasing number of phases warrants further study.

References

Ballard, S., J. Hipp, M. Begnaud, C. Young, A. Encarnacao, E. Chael, and W. S. Phillips (2016a). SALSA3D – A tomographic model of compressional wave slowness in the Earth's mantle for improved travel time prediction and travel time prediction uncertainty, Bull. Seism. Soc. Am. 106, 6, doi:10.1785/0120150271.

Ballard, S., J. Hipp, B. Kraus, A. Encarnacao, and C. Young (2016b). GeoTess: A generalized earth model software utility, Seis. Res. Lett. 87, 3, 719-725, doi:10.1785/0220150222.

Ballard, S., J. Hipp, C. Young, G. T. Barker, and M. Chang (2008). Implementation of a pseudo-bending seismic travel time calculator in a distributed parallel computing environment, in Proc. of the 30th Monitoring Research Review, Portsmouth, VA, 338-346, paper 2-02.

Ballard, S., J. R. Hipp, and C. J. Young (2009). Efficient and accurate calculation of ray theory seismic travel time through variable resolution 3D earth models, Seis. Res. Lett. 80, 6, 989-998, doi:10.1785/gssrl.80.6.989.

Begnaud, M. L. (2005). Using a dedicated location database to enhance the gathering of ground truth information, Los Alamos National Laboratory LA-UR-04-5992, 22 pp.

Begnaud, M. L., D. N. Anderson, S. C. Myers, B. Young, J. R. Hipp, and W. S. Phillips (2020a). Updates to the Regional Seismic Travel Time (RSTT) Model: 2. Path-dependent Travel-time Uncertainty, Pageoph, in review.

Begnaud, M. L., S. C. Myers, B. Young, J. R. Hipp, D. Dodge, and W. S. Phillips (2020b). Updates to the Regional Seismic Travel Time (RSTT) Model: 1. Tomography, Pageoph, in review.

Bondár, I., and K. L. McLaughlin (2009). A new ground truth data set for seismic studies, Seis. Res. Lett. 80, 3, 465-472.

Bondár, I., S. C. Myers, E. R. Engdahl, and E. A. Bergman (2004). Epicentre accuracy based on seismic network criteria, Geophys. J. Int. 156, 483-496.

Bondár, I., and D. Storchak (2011). Improved location procedures at the International Seismological Centre, Geophys. J. Int. 186, 1220-1244, doi:10.1111/j.1365-246X.2011.05107.x.

Firbas, P., K. Fuchs, and W. D. Mooney (1998). Calibration of seismograph network may meet Test Ban Treaty's monitoring needs, Eos Trans. AGU 79, 413-421.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times, Geophys. J. Int. 122, 108-124.

Myers, S., M. Begnaud, S. Ballard, W. Phillips, and M. Pasyanos (2011). Extending regional seismic travel time (RSTT) tomography to new regions, in Proc. of the 2011 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies, Tucson, Arizona, 342-351.

Myers, S. C., M. L. Begnaud, S. Ballard, M. E. Pasyanos, W. S. Phillips, A. L. Ramirez, M. S. Antolik, K. D. Hutchenson, G. S. Wagner, J. J. Dwyer, C. A. Rowe, and D. R. Russell (2010). A crust and upper mantle model of Eurasia and North Africa for Pn travel time calculation, Bull. Seism. Soc. Am. 100, 2, 640-656.

Rowe, C., S. Ballard, M. Begnaud, C. Young, L. Steck, and J. Hipp (2009). Validating 3D geophysical models for use in global travel-time calculation for improved event locations, in Proc. of the 2009 Monitorning Research Review, Tucson, Arizona, 408-415.